



الفلسفة الأخلاقية في عصر الذكاء الاصطناعي في المسؤولية والقيم الإنسانية
دراسة: وصفية

منى محمد عبدالله الشريقي

قسم الفلسفة – كلية الآداب الجميل – جامعة صبراتة

*Moral Philosophy in the Age of Artificial Intelligence: Responsibility and Human Values
A Descriptive Study*

Mona Mohammed Abdullah Al-Shariqi

Department of Philosophy – Faculty of Arts and Fine Arts – Sabratha University

Mona.alshriqi@sabu.edu.ly

تاريخ الاستلام: 2026/06/30 تاريخ المراجعة: 2026/07/20 تاريخ القبول: 2026/7/28- تاريخ النشر: 2026/08/22

الملخص

يتناول هذا البحث التحولات التي أحدثها الذكاء الاصطناعي في مفهوم الفعل الأخلاقي والمسؤولية الإنسانية، ويناقش مدى قدرة التصورات الأخلاقية التقليدية على التعامل مع الأنظمة الخوارزمية والقرارات المؤتمتة. يعتمد البحث المنهج التحليلي النقدي والمقارنة بين أخلاق الواجب والنفعية وأخلاق الفضيلة والمقاربات الحقوقية المعاصرة. ويخلص إلى أن الذكاء الاصطناعي لا يلغي المسؤولية البشرية، بل يعيد توزيعها بين المصمم والمطور والمؤسسة والمستخدم والجهة المنظمة، وأن حماية الكرامة والخصوصية والعدالة تتطلب رقابة بشرية قابلة للمساءلة وشفافية مناسبة للسياق. الكلمات المفتاحية: الفلسفة الأخلاقية؛ الذكاء الاصطناعي؛ المسؤولية؛ الكرامة الإنسانية؛ العدالة الخوارزمية؛ الرقابة البشرية

Abstract

This study examines how artificial intelligence transforms the concepts of moral action and human responsibility. It analyses whether traditional ethical theories can address algorithmic systems and automated decisions. Using a critical analytical and comparative approach, the study discusses deontological, consequentialist, virtue-ethical and rights-based perspectives. It argues that AI does not eliminate human responsibility; rather, it redistributes responsibility among designers, developers, institutions, users and regulators. Human dignity, privacy and fairness require accountable human oversight and context-sensitive transparency.

المقدمة

شهد الفكر الفلسفي عبر مساره التاريخي تحولات مفصلية كلما واجهت البشرية ثورة علمية أو تقنية تمس جوهر الوجود الإنساني؛ فمثلما أعادت الثورة الصناعية صياغة مفاهيم العمل والمجتمع والحرية، وتكفلت الثورة المعرفية بإعادة تشكيل النظرية الوجودية والإبستمولوجية، يفرض عصر الذكاء الاصطناعي (Artificial Intelligence) والأنظمة الذاتية المستقلة تحولاً وجودياً وأخلاقياً هو الأكثر عمقاً وتعقيداً في تاريخ الحضارة الإنسانية.

تتجاوز التحديات الناجمة عن الذكاء الاصطناعي نطاق التداعيات التقنية أو الاقتصادية لتلامس الأسس الفلسفية المعقدة التي بُنيت عليها فلسفة الأخلاق (Moral Philosophy) طوال القرن المنصرمة، ظلت النظرية الأخلاقية— سواء في تقاليد الواجبية الكانطية، أو النفعية الإنجليزيتية، أو أخلاق الفضيلة الأرسطية—ترتكز على افتراض حتمي مفاده: أن الفعل الأخلاقي هو حكر حصري على الكائن البشري، باعتباره الفاعل الأخلاقي (Moral Agent) الوحيد المنفرد بملكات الوعي، القصدية (Intentionality)، الإرادة الحرة، والقدرة على التمييز بين الخير والشر وتوقع التبعات. ومع ظهور الشبكات العصبية الاصطناعية (Artificial Neural Networks)، وتقنيات التعلم العميق (Deep Learning)، والأنظمة الذاتية اتخاذ القرار، انكسر هذا الاحتكار المفهومي؛ إذ أصبحت الآلات تُكلف بتنفيذ مهام إدراكية ومعرفية معقدة كان يُعتقد أنها مقصورة على العقل البشري، بدءاً من التشخيص الطبي، مروراً بالأنظمة القضائية الجنائية والتوظيف، وصولاً إلى الأسلحة المستقلة ذاتية التوجيه. إن انتقال الآلة من دور "الأداة السلبية (Tool)" إلى المطيعة لبيئة المبرمج البشري إلى دور "الوسيط المستقل (Autonomous Agent)" الذي يتخذ قرارات تتطور ذاتياً عبر معالجة كميات مهولة من البيانات، أحدث زلزالاً مفهوماً في البيئة الفلسفية.

تتبدى الإشكالية الكبرى في أن اتخاذ القرار الخوارزمي المستقل غالباً ما يحجب كيفية الوصول إلى النتائج فيما يُعرف في الأدبيات التقنية بـ "الصندوق الأسود (Black Box)"; حيث تعجز معالجات التتبع البرمجي عن شرح المسار العقلاني الذي سلكته الخوارزمية للوصول إلى خيار محدد. يترتب على هذا الشلل الشفافي ظهور ما يُسمى في الفلسفة المعاصرة بـ "فجوة المسؤولية (Responsibility Gap)"; فإذا أدى قرار خوارزمي مستقل إلى كارثة طبية أو تمييز اجتماعي أو خسارة بشرية، يعجز البناء الأخلاقي والقانوني التقليدي عن تحديد "المسؤول الأخلاقي" بوضوح. فالمبرمج لم يكن يتنبأ بالمرجات المعقدة لشبكة التعلم، والمستخدم لا يمتلك المعرفة التقنية للتدخل، والآلة ذاتها تغتر إلى الوعي والذمة الأخلاقية التي تجعلها محلاً للوم أو العقاب.

ومن جهة أخرى، يضع هذا التمدد التقني القيم الإنسانية الجوهرية على المحك. فالإنسان الذي أسست فلسفة الأنوار لمركزيته وكرامته والمنفعة، يجد نفسه في هذا العصر موضوعاً للتقييم والتتميط الإحصائي من قبل أنظمة برمجية تصنفه بناءً على خوارزميات التنبؤ السلوكي، مما يهدد استقلاليته الذاتية وخصوصيته وكرامته. من هنا، تتجاوز الفلسفة الأخلاقية دورها الصوري التقليدي لتصبح ضرورة حيوية لتأطير هندسة التكنولوجيا، وإعادة صياغة الحدود الأخلاقية بين الإنسان والآلة، وضمان ألا تتحول أتمتة الحياة إلى أتمتة للقيم وسلب للحقيقة الإنسانية.

أصبح الذكاء الاصطناعي أحد أهم التحولات التقنية والفكرية في العصر الراهن، لأنه لم يعد مجرد أداة للحساب أو معالجة البيانات، بل صار يشارك في التنبؤ والتصنيف والتوصية واتخاذ القرار. وتظهر تطبيقاته في الطب والتعليم والعمل والإدارة والقضاء والإعلام، الأمر الذي يجعل السؤال الأخلاقي متقدماً على السؤال التقني أحياناً: من المسؤول عندما ينتج النظام قراراً خاطئاً أو تمييزياً؟ وما حدود استقلال الآلة؟ وكيف يمكن حماية الإنسان عندما تصبح بياناته مادة للتنبؤ والتحليل؟

تؤكد اليونسكو أن أخلاقيات الذكاء الاصطناعي ينبغي أن تقوم على حقوق الإنسان والكرامة والعدالة والشفافية والإشراف البشري والاستدامة. كما تشير الأدبيات الفلسفية إلى أن المسألة لا تقتصر على المخاطر التقنية، بل تشمل الخصوصية والاستقلالية والعمل والتفاعل بين الإنسان والآلة والتأثير الاجتماعي الأوسع. ومن هنا يهدف البحث إلى إعادة بناء مفهوم المسؤولية في بيئة خوارزمية لا يعمل فيها الفاعل الأخلاقي دائماً بصورة مباشرة أو منفردة.

أولاً/ الإطار المنهجي :

1/1 مشكلة الدراسة

تتحدد مشكلة الدراسة في الإجابة عن التساؤل الفلسفي المحوري الآتي:

كيف يمكن للفلسفة الأخلاقية المعاصرة إعادة تفكيك وتحديد مفهوم المسؤولية الأخلاقية والحفاظ على أسبقية القيم الإنسانية الجوهرية (كالكرامة والحرية والعدالة) في ظل الفجوة الأخلاقية الناجمة عن تفويض السلطة التقييمية والإنشائية للأنظمة الخوارزمية الذاتية؟

1/2 أهداف الدراسة

تستهدف هذه الدراسة تحقيق الغايات الأكاديمية والفلسفية الآتية:

1. التأسيس المفهومي: إعادة فحص المقولات الأخلاقية الفلسفية الكلاسيكية والمعاصرة في ضوء المتغيرات التقنية الحديثة.
2. تفكيك فجوة المسؤولية: التحليل الفلسفي البنوي لإشكالية "فجوة المسؤولية" الناتجة عن اتخاذ القرارات عبر الخوارزميات غير الشفافة.
3. فحص المهددات القيمية: رصد وتقييم المخاطر التي تفرضها الأتمتة الخوارزمية على القيم الإنسانية الأساسية مثل الاستقلالية الذاتية، الكرامة، والعدالة التوزيعية.
4. بناء أطر حوكمة فلسفية: تقديم مقارنة أخلاقية فلسفية تكفل توجيه هندسة البرمجيات الذكية نحو صون المكتسبات الأخلاقية للبشرية.

1/3 أهمية الدراسة

الأهمية النظرية والمعرفية: تسهم هذه الدراسة في سد النقص الواضح في الأدبيات الفلسفية العربية المعاصرة المتعلقة بـ "أخلاقيات التكنولوجيا" (Ethics of Technology)، وتوفر مادة نظرية تحليلية عميقة تستعرض التقاطعات بين فلسفة العقل، فلسفة الأخلاق، وفلسفة الذكاء الاصطناعي.

الأهمية العملية والتطبيقية: يوفر هذا البحث إطاراً مفهوماً نقدياً للباحثين والمشرعين وصناع السياسات التقنية لمساعدتهم على صياغة موانئ أخلاقية وسياسات تنظيمية تكفل ضبط استخدامات التكنولوجيا المستحدثة دون المساس بالحقوق الإنسانية.

1/4 تساؤلات الدراسة

تسعى الدراسة لإجابة التساؤلات الفرعية الآتية:

- هل تتوافر في أنظمة الذكاء الاصطناعي المستقلة الشروط الفلسفية اللازمة لإسباغ صفة "الفاعل الأخلاقي" عليها؟
- كيف تنشأ "فجوة المسؤولية الأخلاقية" في القرارات الصادرة عن برمجيات التعلم العميق، وكيف يمكن معالجتها فلسفياً؟
- ما هي انعكاسات التنبؤ الخوارزمي والتتقيح السلوكي على القيم الإنسانية الجوهرية كالحرية والكرامة ، الخصوصية؟

- كيف يمكن توظيف النظريات الأخلاقية (كالواجبية والنفعية وأخلاق الفضيلة) لضبط وتأطير تصميم وتطوير تقنيات الذكاء الاصطناعي؟

حدود الدراسة :

1. الحدود الموضوعية: الفلسفة الأخلاقية في عصر الذكاء الاصطناعي: دراسة في المسؤولية والقيم الإنسانية

2. الحدود الزمنية :

1/5 المفاهيم والمصطلحات

الفلسفة الأخلاقية: (Moral Philosophy) الدراسة الفلسفية المعمقة للقيم، والمبادئ، والقواعد التي تميز بين الخير والشر، والعدل والظلم، وتحدد أسس التكليف والمسؤولية عن الأفعال البشريّة.

الذكاء الاصطناعي: (Artificial Intelligence) فرع من علوم الحاسوب يستهدف تطوير أنظمة وبرمجيات قادرة على محاكاة أنماط التفكير الإنساني، والتعلم الآلي من البيانات، واتخاذ قرارات مستقلة حل المعضلات المعقدة. الفاعل الأخلاقي: (Moral Agent) الكيان الفلسفي الذي يمتلك ملكات الإدراك والوعي والإرادة الحرة والتكليف، مما يجعله قادراً على تحمل التبعة الأخلاقية والقانونية عن أفعاله.

فجوة المسؤولية: (Responsibility Gap) المعضلة الفلسفية الأخلاقية التي تنشأ عندما يتسبب نظام تقني مستقل في إحداث ضرر، دون إمكانية إسناد المسؤولية الأخلاقية بوضوح إلى المبرمج، أو المستخدم، أو الآلة ذاتها. القيم الإنسانية: (Human Values) مجموعة المبادئ الوجودية والمواثيق الأخلاقية المطلقة التي تشكل جوهر الكرامة الإنسانية، وتتضمن الحرية، الخصوصية، العدالة، والاستقلال الذاتي.

1/6 منهج الدراسة

تعتمد الدراسة على المنهج الوصفي التحليلي والنقدي؛ وهو المنهج الأكثر ملاءمة للأبحاث الفلسفية النظرية. يقوم المنهج على:

الوصف والتفكير المفاهيمي: مراجعة القواعد الفلسفية والأخلاقية والتصورات التقنية للذكاء الاصطناعي وتفكيك مكوناتها المفهومية.

التحليل النقدي: مقابلة المقولات الفلسفية الكلاسيكية والمعاصرة بالواقع المعرفي لخوارزميات التعلم العميق، وكشف التناقضات والحدود التي تواجه النظرية الأخلاقية عند التعامل مع الآلات المستقلة.

التركيب والحجاج: الاستدلال الفلسفي لبناء رؤية نقدية متماسكة تُعيد تأطير المسؤولية والأخلاق التقنية وفق رؤية إنسانية غائية.

1/7 الدراسات السابقة

1. دراسة ياسين: « (2026) الذكاء الاصطناعي والمسؤولية الأخلاقية»

تناولت الدراسة العلاقة بين تطور أنظمة الذكاء الاصطناعي ومفهوم المسؤولية الأخلاقية، وانطلقت من إشكالية مدى إمكانية اعتبار الذكاء الاصطناعي فاعلاً أخلاقياً مستقلاً. واعتمدت المنهج الفلسفي التحليلي، مع استعراض تصورات أرسطو وكانط وجون ستيوارت ميل وسارتر حول الحرية والإرادة والوعي والمسؤولية. وتوصلت إلى أن أنظمة الذكاء الاصطناعي لا تتوافر فيها بصورة كاملة شروط الفاعل الأخلاقي، ولذلك ينبغي توزيع المسؤولية بين المطورين

والمستخدمين والمؤسسات والجهات التنظيمية، مع ضرورة وضع أطر أخلاقية وقانونية تضمن الشفافية والمساءلة وتحمي القيم الإنسانية.

2.دراسة» الذكاء الاصطناعي بين الفلسفة والأخلاق وتحديات المستقبل(2026) «

بحثت الدراسة في الأسس الفلسفية للذكاء الاصطناعي والتحديات الأخلاقية المرتبطة باستخداماته المتزايدة في الحياة الإنسانية. وركزت على أثر الذكاء الاصطناعي في التربية الأخلاقية والقيم الإنسانية والمخاطر الناتجة عن الاستخدام غير المسؤول للتكنولوجيا. وانتهت إلى أهمية توجيه تطبيقات الذكاء الاصطناعي بما يخدم الإنسان والمجتمع، وتحقيق التوازن بين الاستفادة من التطور التقني والمحافظة على القيم الإنسانية.

3.دراسة تيرس»: (2022) أزمة الإنسان المعاصر بين الواقع التكنولوجي والتوجيه الأخلاقي في فلسفة إريك

فروم»

ركزت الدراسة على أزمة الإنسان المعاصر في ظل التطور العلمي والتكنولوجي من خلال تحليل فلسفة إريك فروم الإنسانية. وبيّنت أن التطور التكنولوجي لا يؤدي بالضرورة إلى تحقيق السعادة الإنسانية، بل قد يزيد الاغتراب وفقدان المعنى إذا انفصل عن التوجيه الأخلاقي. وأكدت أهمية بناء أخلاق إنسانية للعلم والتكنولوجيا تجعل التقدم التقني وسيلة لخدمة الإنسان.

الدراسات الأجنبية

الدراسة الأجنبية الأولى(2019) Floridi & Cows

هدفت الدراسة إلى تحليل مجموعة من المبادئ الأخلاقية المقترحة للذكاء الاصطناعي، وخلصت إلى وجود تقارب حول مبادئ المنفعة وعدم الإضرار والاستقلالية والعدالة، مع إضافة مبدأ قابلية التفسير الذي يرتبط بفهم كيفية عمل النظام وبمسألة تحديد المسؤولية. وتفيد الدراسة الحالية في بناء الإطار الأخلاقي للمسؤولية والقيم الإنسانية.

الدراسة الأجنبية الثانية(2019) Jobin, Ienca & Vayena :

حللت الدراسة المبادئ والإرشادات العالمية لأخلاقيات الذكاء الاصطناعي، ووجدت تقارباً حول الشفافية والعدالة وعدم الإضرار والمسؤولية والخصوصية، مع اختلاف في تفسير هذه المبادئ وطرق تطبيقها. وتفيد هذه النتيجة في تأكيد أن أخلاقيات الذكاء الاصطناعي تحتاج إلى تحويل المبادئ العامة إلى آليات عملية للحوكمة والمساءلة.

التعقيب على الدراسات السابقة

توضح الدراسات السابقة أن الأخلاق والذكاء الاصطناعي أصبحا من القضايا الفلسفية المعاصرة التي تستدعي إعادة النظر في المسؤولية والحرية والإرادة والقيم. وتتميز الدراسة الحالية بمحاولة الجمع بين المسؤولية الأخلاقية والقيم الإنسانية ومكانة الإنسان في عصر الذكاء الاصطناعي، مع الاستفادة من التصورات الفلسفية الكلاسيكية والمعاصرة.

ثانيا/الإطار النظري

تمهيد

يمثل الذكاء الاصطناعي إحدى أبرز التحولات التقنية والفكرية في العصر الحديث، إذ لم تعد تطبيقاته تقتصر على المجالات الصناعية والحسابية، بل امتدت إلى التعليم والصحة والإدارة والإعلام والاقتصاد واتخاذ القرار. وقد أدى هذا الانتشار إلى ظهور أسئلة فلسفية جديدة حول طبيعة العلاقة بين الإنسان والآلة، وحدود المسؤولية عن القرارات الآلية، ومكانة القيم الإنسانية في بيئة تتزايد فيها قدرة الأنظمة الذكية على التحليل والتنبؤ والتوليد واتخاذ التوصيات. ومن هنا أصبح من الضروري أن يواكب التطور التقني تطوراً أخلاقياً وفلسفياً يحدد ما ينبغي فعله، وليس فقط ما يمكن فعله.

أصبح الذكاء الاصطناعي جزءاً من البنية اليومية للمجتمعات الحديثة، بعد أن انتقل من نطاق المختبرات المتخصصة إلى مجالات التعليم والصحة والإدارة والاقتصاد والإعلام والأمن. ولا ترجع أهمية الموضوع إلى قدرة الخوارزميات على معالجة البيانات أو إنتاج التنبؤات فحسب، بل إلى كونها تشارك بصورة مباشرة أو غير مباشرة في قرارات تمس فرص الأفراد وخصوصيتهم وحريتهم وكرامتهم. ومن ثم فإن السؤال الأخلاقي لا يقتصر على ما إذا كان النظام يعمل بكفاءة، وإنما يمتد إلى الغاية التي صُمم من أجلها، والآثار التي يتركها، والجهة التي تتحمل مسؤولية نتائجه (UNESCO). ينطلق هذا الإطار من فكرة مركزية مؤداها أن الذكاء الاصطناعي ليس فاعلاً مستقلاً عن المجتمع، بل هو نتاج قرارات بشرية ومؤسسية تبدأ من اختيار المشكلة والبيانات والنموذج، وتستمر في البرمجة والتدريب والاختبار والنشر والمراقبة. لذلك لا ينبغي أن تُفهم أخلاقيات الذكاء الاصطناعي بوصفها مجموعة شعارات عامة، بل بوصفها مجالاً يربط المبادئ الفلسفية بالتصميم التقني والقواعد القانونية وبالممارسات المؤسسية. وقد أظهرت الأدبيات الدولية تقارباً ملحوظاً حول الشفافية والعدالة وعدم الإضرار والمسؤولية والخصوصية، مع استمرار الاختلاف في تفسير هذه المبادئ وترتيب أولوياتها وآليات تنفيذها.

مفهوم الفلسفة الأخلاقية

الفلسفة الأخلاقية فرع من فروع الفلسفة يبحث في طبيعة الخير والشر والصواب والخطأ والواجب والفضيلة والمسؤولية. وهي لا تكتفي بوصف السلوك الإنساني، وإنما تحاول تقديم معايير عقلية للحكم عليه. وتكتسب الأخلاق أهميتها في عصر الذكاء الاصطناعي لأن القرارات التقنية قد تؤثر بصورة مباشرة في حقوق الأفراد وحررياتهم وخصوصيتهم وفرصهم في التعليم والعمل والخدمات.

الفلسفة الأخلاقية هي البحث المنظم في طبيعة الخير والشر، والصواب والخطأ، والواجب والفضيلة، والمعايير التي ينبغي أن توجه الفعل الإنساني. وهي لا تكتفي بوصف السلوك كما يقع، بل تسأل عن الأساس الذي يجعل فعلاً ما جيداً بالقبول أو الرفض. وتستند المقاربات الأخلاقية إلى أسئلة معيارية مثل: ما الذي ينبغي فعله؟ وما حدود الحرية؟ ومتى يكون الفاعل مسؤولاً؟ وكيف نوفق بين مصلحة الفرد ومصلحة الجماعة؟ وتزداد قيمة هذه الأسئلة في مجال الذكاء الاصطناعي لأن الأنظمة الذكية قد تؤثر في الإنسان دون أن تمتلك فهماً إنسانياً للمعنى أو القدرة على تحمل اللوم بالمعنى الأخلاقي الكامل.

وتتخذ الفلسفة الأخلاقية في هذا البحث وظيفة تحليلية ونقدية. فهي تحلل المفاهيم التي تُستخدم في النقاش التقني مثل الاستقلالية والعدالة والخصوصية والثقة، وتختبر ما إذا كان استخدامها متسقاً مع الواقع الاجتماعي. كما تنقد الافتراضات القائلة إن زيادة الدقة التقنية تعني بالضرورة زيادة الخير الأخلاقي؛ فقد يكون النظام دقيقاً في التنبؤ لكنه يكرس تمييزاً تاريخياً، أو فعلاً في المراقبة لكنه ينتهك الخصوصية، أو سريعاً في اتخاذ القرار لكنه يحرم الشخص من فرصة الاعتراض. لذلك تتطلب الأخلاق تقييماً للغايات والوسائل والنتائج معاً

مفهوم الذكاء الاصطناعي من المنظور الفلسفي

يمكن النظر إلى الذكاء الاصطناعي فلسفياً باعتباره محاولة لبناء أنظمة قادرة على تنفيذ مهام ارتبطت تقليدياً بالقدرات العقلية للإنسان، مثل التعلم والاستدلال والتنبؤ وفهم اللغة والتعرف على الأنماط. ولا يقتصر السؤال الفلسفي على قدرة الآلة، بل يتجاوز ذلك إلى سؤال طبيعة الذكاء نفسه: هل الذكاء مجرد معالجة للمعلومات، أم أنه يتضمن الوعي والقصد والخبرة الذاتية؟ وهذا التمييز مهم عند مناقشة المسؤولية الأخلاقية.

يقصد بالذكاء الاصطناعي، في الإطار العملي، الأنظمة القائمة على الآلات التي تستنتج من المدخلات كيفية توليد مخرجات مثل التنبؤات أو التوصيات أو القرارات أو المحتوى، بما يؤثر في بيئات مادية أو افتراضية. وتعمل هذه الأنظمة عبر دورة حياة تبدأ بجمع البيانات وتصميم النموذج وتدريبه، ثم اختباره ونشره وتحديثه أو إيقافه. ويكشف هذا التعريف أن المسؤولية لا ترتبط بلحظة تشغيل النظام فقط؛ إذ يمكن أن ينشأ الضرر من بيانات غير ممثلة، أو من معيار تصميم غير مناسب، أو من نشر النظام في سياق لم يُختبر فيه، أو من غياب المراقبة بعد النشر (OECD). وتتميز أنظمة الذكاء الاصطناعي بعدة خصائص تجعلها موضوعاً أخلاقياً خاصاً. أولها التعقيد، إذ قد يصعب تتبع العلاقة بين المدخلات والمخرجات. وثانيها الاعتماد على البيانات، بما يجعل الانحيازات الاجتماعية قابلة للانتقال إلى النموذج. وثالثها قابلية التوسع، حيث يمكن لقرار واحد أو نموذج واحد أن يؤثر في أعداد كبيرة من الأشخاص. ورابعها التكيف والتحديث، الذي قد يغير سلوك النظام بمرور الوقت. وخامسها قابلية التوزيع المؤسسي للمسؤولية بين المطور والمشغل والمستخدم ومزود البيانات والجهة المنظمة. ولهذا ركزت الأطر الحديثة على التتبع والتوثيق وتحديد الأدوار والمراجعة البشرية طوال دورة الحياة.

الأسس الفلسفية لأخلاقيات الذكاء الاصطناعي

المنظور النفعي ومنع الضرر وتعظيم المنفعة

يركز المنظور النفعي على نتائج الفعل، ويقيم القرار بحسب مقدار المنفعة أو الرفاه الذي يحققه مقارنة بالضرر الذي قد ينتج عنه. وفي مجال الذكاء الاصطناعي يمكن أن يدعم هذا المنظور استخدام الأنظمة في التشخيص الطبي أو إدارة المرور أو اكتشاف الاحتيال إذا كانت النتائج المتوقعة تحسن حياة الناس. غير أن الحساب النفعي يواجه صعوبة أخلاقية عندما تُوزع المنافع والأضرار بصورة غير متساوية؛ فقد يحقق النظام منفعة عامة ظاهرة لكنه يحمل فئة ضعيفة تكلفة أكبر. لذلك لا يكفي قياس الناتج الإجمالي، بل ينبغي فحص من يستفيد ومن يتضرر وكيف يمكن تعويض المتضررين. ويترجم مبدأ عدم الإضرار هذا القلق إلى واجب عملي، مثل اختبار النظام قبل نشره، وتحديد الاستخدامات غير المقبولة، ووضع آليات توقف آمنة، ومراقبة النتائج غير المقصودة. وتؤكد الأطر الحديثة أن النظام ينبغي أن يكون آمناً وقابلاً للتجاوز أو الإصلاح أو الإيقاف عندما يظهر سلوك ضار. لكن عدم الإضرار لا يعني تجسيد الابتكار؛ بل يعني إدارة المخاطر بصورة متناسبة مع جسامته والضرر واحتماله، مع إيلاء عناية خاصة للأنظمة التي تمس الحقوق الأساسية (OECD, 2019, Section 1.4)؛ (NIST, 2023, pp. 20–22).

المنظور الواجبي وكرامة الإنسان

ينظر المنظور الواجبي إلى الأخلاق باعتبارها التزاماً بمبادئ وحقوق لا يجوز انتهاكها حتى لو أمكن تبرير الانتهاك بنتيجة نافعة. ومن تطبيقاته في الذكاء الاصطناعي أن لا يُعامل الفرد بوصفه وسيلة لتحقيق هدف إحصائي أو تجاري، وأن تحترم خصوصيته وحقه في عدم التمييز وحقه في معرفة أن قراراً آلياً أثر فيه. ويظهر هذا المنظور بوضوح في جعل حقوق الإنسان والكرامة الإنسانية أساساً لأي تنظيم للذكاء الاصطناعي، وفي رفض تحويل الإنسان إلى رقم أو درجة أو تصنيف دون ضمانات (UNESCO, 2022, pp. 1–2)؛ (OECD, 2019, Section 1.2). وتتطلب كرامة الإنسان ألا تُسند القرارات الحساسة إلى النظام بصورة تلغي الحكم البشري أو تجعل الاعتراض شكلياً. فالإنسان لا يحتاج فقط إلى نتيجة صحيحة، بل إلى معاملة عادلة وإمكانية فهم القرار والطعن فيه. ومن ثم تتكامل

الكرامة مع الشفافية والقابلية للتفسير والإنصاف الإجرائي، لأن غياب هذه العناصر قد يحول النظام إلى سلطة غير مرئية يصعب مساءلتها) .

أخلاق الفضيلة والمسؤولية المهنية

تركز أخلاق الفضيلة على صفات الفاعل الأخلاقية، مثل النزاهة والحكمة والشجاعة والاعتدال والحرص على الصالح العام. وتفيد هذه المقاربة في الذكاء الاصطناعي لأنها تنقل النقاش من سؤال: هل التزم المطور بقائمة متطلبات؟ إلى سؤال: هل تصرف المطور أو المؤسسة بنزاهة ومسؤولية؟ فقد يكون الفعل مطابقاً شكلياً للسياسة لكنه يفتقر إلى الشفافية أو يتجاهل آثاراً متوقعة على الفئات الضعيفة. لذلك تحتاج المؤسسات إلى ثقافة أخلاقية ومهنية تشجع الإبلاغ عن المخاطر، وتكافئ التصحيح المبكر، وتمنع إخفاء عيوب النظام (Jobin, Ienca & Vayena, 2019, pp. 15-16؛ NIST, 2023, pp. 21-22).

العدالة والرعاية والتضامن

تعني العدالة في الذكاء الاصطناعي أكثر من مساواة شكلية في المعاملة؛ فهي تشمل عدالة البيانات، وعدالة النموذج، وعدالة النتائج، وعدالة الإجراءات. وقد تكون الخوارزمية متساوية في تطبيق قاعدة واحدة، لكنها تنتج أثراً غير عادل بسبب اختلاف الظروف التاريخية أو الاجتماعية بين الفئات. لذلك يجب فحص التمثيل في البيانات، ومعدلات الخطأ بين المجموعات، وشروط الوصول إلى الخدمة، وإمكان الاعتراض والتعويض. وترتبط الأدبيات بين العدالة والشفافية والمسؤولية، لأن كشف القرار وتحديد صاحبه شرطان لمراجعة أثره.

أما أخلاق الرعاية والتضامن فتضيفان بعداً يركز على العلاقات والاعتماد المتبادل وحماية الفئات الأكثر هشاشة. فالقرار الآلي لا يقع على أفراد منعزلين دائماً، بل داخل مؤسسات وعلاقات اجتماعية قد تجعل بعض الأشخاص أقل قدرة على فهم النظام أو الاعتراض عليه. ويبرز التضامن في توزيع فوائد الذكاء الاصطناعي ومخاطره، وفي ضمان ألا تؤدي الأتمتة إلى تعميق الإقصاء أو تفكيك الحماية الاجتماعية. وقد ربطت الدراسات العالمية بين أخلاقيات الذكاء الاصطناعي وقضايا العمل والتماسك الاجتماعي وحماية الفئات الضعيفة ..

المسؤولية الأخلاقية في عصر الذكاء الاصطناعي

المسؤولية الأخلاقية تعني قابلية نسبة الفعل أو القرار إلى فاعل يمكن مساءلته عن علمه واختياره وقدرته على التوجيه والتصحيح. وعلى الرغم من استخدام تعبيرات مثل "النظام المستقل" أو "الوكيل الذكي"، فإن الاستقلال الوظيفي لا يساوي بالضرورة استقلالاً أخلاقياً. فالآلة قد تنفذ أهدافاً أو تتعلم أنماطاً، لكنها لا تمتلك بالضرورة وعياً أخلاقياً أو قدرة على فهم قيمة الضرر والعدل أو قبول اللوم والاعتذار. ولهذا أكدت اليونسكو أن المسؤولية النهائية والمساءلة يجب أن تبقى على عاتق أشخاص طبيعيين أو اعتباريين، وألا تمنح الأنظمة الذكية شخصية قانونية لمجرد قدرتها على اتخاذ مخرجات. وتنشأ مشكلة "فجوة المسؤولية" عندما يؤدي توزيع الأدوار وتعقيد النموذج إلى صعوبة تحديد الشخص أو الجهة التي ينبغي مساءلتها. ولا تعالج هذه الفجوة بإسناد الخطأ إلى الخوارزمية، لأن ذلك قد يحجب القرارات البشرية التي سبقت تشغيلها. والمعالجة الأقوى هي بناء سلسلة مسؤولية قابلة للتتبع تشمل مالك النظام، ومصمم النموذج، ومزود البيانات، والجهة التي نشرته، والمستخدم المهني، والجهة الرقابية. ويعني ذلك أن المسؤولية قد تكون موزعة، لكنها ليست مجهولة؛ إذ ينبغي تحديد الصلاحيات والواجبات ونقاط المراجعة مسبقاً .

المسؤولية عبر دورة حياة النظام : تبدأ المسؤولية الأخلاقية قبل كتابة الشيفرة، عند تحديد المشكلة والغاية والفئة المستهدفة. ففي هذه المرحلة يجب سؤال :هل المشكلة مناسبة للحل الآلي؟ وهل توجد مخاطر على الحقوق؟ وهل توجد فئات قد تتأثر دون قدرة على المشاركة؟ ثم تمتد المسؤولية إلى مرحلة البيانات، حيث يجب التحقق من المصدر والتمثيل والجودة والخصوصية. وفي مرحلة التصميم والتدريب ينبغي اختبار الأداء والانحياز والأمن. وبعد النشر يجب مراقبة الاستخدام الفعلي وتسجيل الحوادث وإتاحة التصحيح والإيقاف..ويقترح إطار NIST إدارة المخاطر بصورة مستمرة عبر أربع وظائف مترابطة :الحوكمة، وتحديد السياق والمخاطر، والقياس، والإدارة. وتكمن أهمية هذا التصور في أنه لا يعامل الأخلاق كفحص نهائي، بل كعملية مستمرة تتكرر مع تغير البيانات والسياق والأهداف. كما يربط الحوكمة بالمبادئ المؤسسية، ويوصي بتوثيق الأدوار والقرارات ونتائج المراجعة، بما يدعم الشفافية والمراجعة البشرية والمساءلة.

الشفافية والتفسير والمساءلة: الشفافية لا تعني دائماً الكشف الكامل عن الشيفرة، بل توفير معلومات مفهومة ومناسبة عن غرض النظام، وحدوده، ومصادر بياناته، وطريقة استخدام مخرجاته، والجهات المسؤولة عنه. أما القابلية للتفسير فتعني تمكين الشخص المتأثر من فهم الأسباب ذات الصلة بالنتيجة، خصوصاً عندما يترتب عليها رفض طلب أو حرمان من فرصة أو فرض قيد. وقد ربط إطار Al4People بين قابلية الفهم وبين المساءلة، لأن سؤال "كيف صدر القرار؟" يتصل بسؤال "من المسؤول عن طريقة صدوره؟". "ومع ذلك، يجب الحذر من جعل التفسير إجراءً شكلياً. فالتفسير المفيد ينبغي أن يكون دقيقاً ومفهوماً ومرتبئاً بإمكان الاعتراض، لا مجرد وصف تقني لا يستطيع المتضرر استعماله. كما ينبغي أن يراعي التوازن بين الشفافية وحماية الخصوصية والأسرار التجارية والأمن. ولهذا تبرز الحاجة إلى سجلات القرار، وتوثيق نسخ النموذج والبيانات، وتحديد مسارات الشكوى والتعويض، وإجراء تدقيق مستقل عندما يكون أثر النظام مرتفعاً ..

العدالة وعدم التحيز الخوارزمي

لا تكون الخوارزمية محايدة بالضرورة لمجرد أنها رياضية. فالبيانات التي تتعلم منها قد تحمل تحيزات تاريخية أو اجتماعية، كما يمكن أن تؤدي معايير اختيار البيانات أو تصميم النموذج إلى نتائج غير متوازنة. ومن منظور فلسفة العدالة، يجب فحص آثار النظام على الفئات المختلفة، والتأكد من عدم تحويل الاختلافات الاجتماعية القائمة إلى صور جديدة من التمييز الرقمي.

الخصوصية والكرامة الإنسانية

يتيح الذكاء الاصطناعي معالجة كميات هائلة من البيانات الشخصية، الأمر الذي يجعل الخصوصية من أهم القيم التي ينبغي حمايتها. والخصوصية ليست مجرد إخفاء للمعلومات، بل ترتبط باستقلال الفرد وقدرته على التحكم في جوانب من حياته. ومن ثم فإن استخدام البيانات يجب أن يخضع للضرورة والتناسب والشفافية والموافقة والضمانات القانونية.

الشفافية وقابلية التفسير

كلما ازدادت تعقيدات النماذج الذكية، أصبح من الصعب على المستخدم معرفة كيفية الوصول إلى النتيجة. ويثير ذلك مشكلة أخلاقية عندما تكون النتيجة مؤثرة في حياة الإنسان. لذلك تبرز أهمية الشفافية، وإتاحة تفسير مناسب للقرارات، ووجود آليات للمراجعة والاعتراض والتصحيح.

الإنسان في مواجهة التحول التكنولوجي

لا ينبغي أن يكون الهدف من التطور التقني استبدال الإنسان في كل مجال، وإنما دعم قدراته وتحسين جودة حياته. ويقتضي ذلك فلسفة إنسانية للتقنية تجعل الإنسان غاية أساسية، وتضع حدوداً واضحة لاستخدام القوة التقنية. كما ينبغي تطوير التربية الأخلاقية الرقمية حتى يصبح المستخدم قادراً على فهم مخاطر التقنية ومسؤوليته تجاه الآخرين. تفكيك مفهوم "الفاعل الأخلاقي" وإشكالية الصندوق الأسود

ترتكز النظرية الأخلاقية الكلاسيكية، ولا سيما في تقاليد الكانطية، على الامتثال لمعايير صارمة تُعرف بالحرية والإرادة الواعية والقصدية. فالطرح الكانطي يقرن التكليف الأخلاقي بالاستقلال الذاتي (Autonomy) للذات العاقلة التي تمتلك الوعي بالشرط الإنساني والتنبؤ المسبق بالغايات والنتائج. ومع استبدال الآليات البرمجية التقليدية ببرمجيات الشبكات العصبية المتقدمة وتقنيات التعلم العميق، برز نمط غير مسبوق من القرارات التي تنقر إلى إمكانية الشرح والتعليل الإنساني.

تتضح معضلة "الصندوق الأسود" (Black Box) "في كون الخوارزميات المتقدمة تعالج ملايين المتغيرات والبيانات الضخمة عبر أوزان ومعاملات رياضية معقدة تتغير ذاتياً أثناء عملية التعلم، مما يجعل تتبع المسار التحليلي المنطقي للقرار الذي اتخذته الآلة أمراً ممتنعاً حتى على المبرمجين المصممين للنظام نفسه. يُنتج ذلك انفصلاً كلياً بين القصدية البشرية الشارحة وبين المخرج البرمجي العملي.

إن الإصرار على إسباغ صفة "الفاعل الأخلاقي" (Moral Agent) "على الآلة يُعد خطأً فلسفياً وإبستمولوجياً خطيراً بين شقين مختلفين تماماً:

الأول: المعالجة الحسابية الإحصائية: (Computational Processing) وهي قدرة فائقة على الفرز والترشيح والتنبؤ بناءً على أنماط رياضية مجردة تخلو من العاطفة والوعي.

الثاني: الوعي الأخلاقي والقصدية الذاتية: (Intentional Consciousness) وهو التجربة الباطنية الواعية التي تدرك المعنى الأخلاقي للواجب، وتتألم للخطأ، وتحمل الملامة والتبعة الأخلاقية.

بناءً على ذلك، تنتهي التحليلات الفلسفية النقدية إلى أن الآلة—مهما بلغت مستويات استقلالها الذاتي البرمجي—تظل مجرد "فاعل تقني" (Technical Agent) "سلبياً، وتجريدها من الوعي والذمة الأخلاقية يمنع منطقياً وإبستمولوجياً تحميلها المساءلة الأخلاقية، مما يُبقي الالتزام الأخلاقي حكرًا حصرياً على العنصر البشري المصمم، والمطور، والمشغل.

"فجوة المسؤولية" والانعكاسات على النظريات الأخلاقية الكلاسيكية : تتجسد "فجوة المسؤولية" (Responsibility Gap) بصفقتها الأزمة الأخلاقية الأكثر إلحاحاً في التفكير الفلسفي المعاصر؛ حيث تنشأ هذه الفجوة في المساحة الضبابية الفاصلة بين الآلة ذاتية القرار والأنظمة البشرية التي تطورها. يمكن إبراز حدود هذه المعضلة بمواجهة النظريات الأخلاقية الكلاسيكية الممتدة:

الأخلاق الواجبية: (Deontological Ethics)

تعتمد فلسفة الواجب عند إيمانويل كانط على الإرادة الطيبة والامتثال للقانون الأخلاقي كغاية بحد ذاته. تتهار هذه المقاربة أمام تقنيات الذكاء الاصطناعي نتيجة غياب "الإرادة" أصلاً لدى البرمجيات؛ فالآلة لا تلتزم بالقواعد الواجبة من منطق المنفعة والوعي بالواجب، بل تنفذاً لتعليمات برمجية مجردة من الغائية الأخلاقية الذاتية.

الأخلاق النفعية: (Utilitarianism)

تُقاس الأخلاق عند جيريمي بينثام وجون ستيوارت ميل بالنتائج المترتبة على الفعل وتحقيق السعادة الكبرى لأكبر عدد ممكن. ورغم أن الخوارزميات تُصمم في كثير من الأحيان لتحسين الكفاءة والمخرجات النفعية، إلا أن عجز الآلة عن استيعاب القيم الإنسانية الكيفية (كالعدالة والكرامة) وتفضيلها للأرقام والكميات المجردة يؤدي إلى نتائج كارثية؛ مثل التضحية بأفراد أو فئات مستضعفة في سبيل معايير تحسين رياضية جافة.

أخلاق الفضيلة: (Virtue Ethics)

ترتكز فلسفة الفضيلة عند أرسطو على بناء الشخصية الأخلاقية الفاضلة والممارسة العملية والتمرس الاجتماعي على العقلانية والمقاصد السامية. تتلشى هذه الفلسفة تماماً عند تطبيقها مباشرة على الآلة؛ فالآلة لا تمتلك مساراً حياتياً أو تجربة وجودية تمكنها من اكتساب الفضائل أو صقل حكمتها العملية. (Phronesis) تُظهر هذه التحليلات أن النظريات الأخلاقية الكلاسيكية لا يمكن إسقاطها بشكل آلي ومباشر على الأنظمة البرمجية المستقلة، بل يجب تحويل هذه النظريات إلى أطر وموجهات تُفرض نقدياً على الفاعل البشري (المبرمج والمؤسسات) في مراحل التخطيط، الهندسة، والنشر.

مهددات القيم الإنسانية (الكرامة، الاستقلالية، والعدالة)

لا تقتصر آثار الذكاء الاصطناعي على تحديات المساواة، بل تمتد لتشكل تهديداً هكلياً ومباشراً للقيم الإنسانية الجوهرية التي أسست لها الفلسفة الحديثة:

تآكل الكرامة والاستقلال الذاتي: (Autonomy)

تحول الإنسان في ظل أنظمة التنبؤ السلوكي والتنفيذ الخوارزمي من "ذات حرة أصيلة" إلى مجرد "كائن رقمي إحصائي" يُصنف ويُحدد اختياراته السياسية والاستهلاكية بناءً على بياناته السابقة. إن ممارسات التوجيه الخوارزمي الناعم تُجرد الإنسان من ممارسة حريته الفردية الحقيقية وتفرغ قراراته من العفوية والأصالة.

تضخيم الظلم الاجتماعي ومأسسة التمييز:

تعتمد تقنيات التعلم الآلي على التغذية ببيانات تاريخية ضخمة. ونظراً لأن هذه البيانات المجمعّة تعكس واقعاً إنسانياً محمل بالتحيزات العرقية، الطبقية، والنوعية، فإن الخوارزميات تعيد إنتاج هذه التحيزات وتضخمها وتحولها إلى قرارات آلة مجردة تُضفي على التمييز صفة "الحياد الرقمي".

انتهاك الخصوصية والسيطرة المظلمة :

تتطلب الخوارزميات المتقدمة مراقبة مستمرة وجمعاً شاملاً للبيانات الفردية والاجتماعية. تؤدي هذه المراقبة الدائمة إلى استلاب الخصوصية الفردية، وتحويل الإنسان إلى موضوع مكشوف ومستباح أمام القوى التكنولوجية والتجارية، مما يقوض القاعدة الفلسفية التي تعتبر الخصوصية شرطاً أساسياً للحرية والتفكير المستقل.

القيم الإنسانية الحاكمة للذكاء الاصطناعي :

الكرامة والاستقلالية والحرية : تتطلب الكرامة ألا يُختزل الإنسان في خصائص قابلة للقياس أو في احتمالات إحصائية. أما الاستقلالية فتشير إلى قدرة الفرد على تكوين اختياراته وفهمها والتصرف وفقها دون تلاعب أو إكراه خفي. وقد يدعم الذكاء الاصطناعي الاستقلالية عندما يقدم معلومات وأدوات مساعدة، لكنه قد يضعفها عندما يستخدم التخصيص للتلاعب أو عندما يفرض توصية لا يمكن فهمها أو رفضها. لذلك توصي المبادئ الدولية بالحفاظ على الوكالة البشرية والرقابة البشرية والقدرة على التدخل في السياقات المناسبة.

لخصوصية وحماية البيانات : تتجاوز الخصوصية في عصر الذكاء الاصطناعي فكرة سرية المعلومات إلى التحكم في كيفية جمع البيانات وربطها وتحليلها واستنتاج معلومات جديدة منها .فقد يستنتج النظام حالات صحية أو ميولاً أو علاقات اجتماعية لم يصرح الفرد بها صراحة .ولذلك تتطلب الحماية تحديد الغرض، وتقليل البيانات، وحماية البيانات الحساسة، وإعلام الأفراد بكيفية الاستخدام، وتوفير إمكان الوصول والتصحيح والحذف حيث يسمح القانون .كما ينبغي منع الاستخدامات التي تنتهك الكرامة أو تحول المراقبة إلى تصنيف اجتماعي واسع.

العدالة وعدم التمييز والشمول : ينشأ التمييز الخوارزمي عندما تؤدي البيانات أو المعايير أو طريقة الاستخدام إلى معاملة غير عادلة لفئة معينة .وقد يكون التمييز مباشراً أو غير مباشر، معلناً أو ناتجاً عن متغيرات بديلة ترتبط بالعرق أو النوع أو الوضع الاجتماعي أو الموقع الجغرافي .وتحتاج العدالة إلى مشاركة الفئات المتأثرة، ومراجعة جودة البيانات، واختبار النتائج بين المجموعات، وتفسير الفروق، وإتاحة الاعتراض .كما أن الشمول لا يقتصر على إتاحة المنتج، بل يشمل إشراك جهات نظر متنوعة في التصميم والتقييم والحوكمة.

المنفعة والاستدامة والسلامة : تتصل المنفعة بقدرة الذكاء الاصطناعي على تحسين الصحة والتعليم والخدمات والإنتاجية، بينما تتصل الاستدامة بآثاره على البيئة والموارد والعمل والتماسك الاجتماعي .ولا يجوز تقييم المنفعة بمعزل عن تكلفة التدريب والتشغيل واستهلاك الطاقة أو عن آثار الأتمتة على العاملين .لذلك تجمع الأطر الحديثة بين الرفاه الإنساني والتنمية المستدامة وحماية البيئة، وتدعو إلى توجيه الابتكار نحو منافع مشتركة وتقليل أوجه عدم المساواة..

الحوكمة الأخلاقية وبناء الذكاء الاصطناعي المسؤول

الحوكمة الأخلاقية هي مجموعة القواعد والعمليات والهياكل التي تضمن ترجمة القيم إلى قرارات قابلة للتنفيذ والمراجعة . وهي تشمل تحديد المسؤوليات، وتقييم الأثر، والتدقيق، وإدارة المخاطر، ومشاركة أصحاب المصلحة، والتوثيق، وآليات الشكوى والتعويض .وتكمن وظيفتها في منع انفصال القرار التقني عن القرار الأخلاقي؛ إذ يجب أن تعرف المؤسسة لماذا استخدمت النظام، وما المخاطر المقبولة، ومن يملك صلاحية الإيقاف أو التصحيح .ويقوم الذكاء الاصطناعي المسؤول على مجموعة ممارسات مترابطة :تقييم أخلاقي قبل النشر، وتصميم يراعي الحقوق، وتوثيق مصادر البيانات وحدود النموذج، واختبارات للانحياز والأمن، ومراقبة بعد النشر، وتدريب للعاملين، ومراجعة مستقلة في الأنظمة عالية الأثر .كما ينبغي إشراك المتخصصين في القانون والأخلاق والمجتمع المدني والفئات المتأثرة، لأن التقييم التقني وحده لا يحدد دائماً ما إذا كان الاستخدام مقبولاً اجتماعياً

أهم النتائج:

- المسؤولية الأخلاقية عن استخدام الذكاء الاصطناعي لا يمكن نقلها إلى الآلة، بل تظل مرتبطة بالأفراد والمؤسسات المسؤولة عن تصميمها وتشغيلها واستخدامها .
- أصبحت قضايا الحرية والاستقلالية والخصوصية والكرامة الإنسانية أكثر تعقيداً مع انتشار الأنظمة الخوارزمية.
- الشفافية والمساءلة ضروريان للحد من القرارات الخوارزمية غير العادلة أو المتحيزة.
- يجب أن تبقى الرقابة البشرية عنصراً أساسياً في القرارات التي تمس حقوق الإنسان وحياته.
- لا ينبغي النظر إلى التقدم التقني باعتباره غاية مستقلة، بل يجب أن يكون مرتبطاً بالقيم الإنسانية والعدالة والكرامة.

- المؤسسات التعليمية والثقافية لها دور مهم في نشر التفكير النقدي والوعي الأخلاقي عند التعامل مع التقنيات الحديثة.

التوصيات

1. صياغة ميثاق أخلاقي عربي للذكاء الاصطناعي: إعداد إطار فلسفي وأخلاقي ملزم ينظم أبحاث هندسة البرمجيات، ويضمن حظر استخدام الخوارزميات غير الشفافة في القضايا التي تمس المصير والحقوق الإنسانية.
2. إقرار المبدأ الإنساني في القرار: (Human-in-the-Loop) التشديد التشريعي والفلسفي على ضرورة الإبقاء على سلطة التقييم والقرار النهائي بيد الفاعل البشري في شتى التطبيقات الحساسة (كالطب، القضاء، والأسلحة).
3. إدماج الفلسفة الأخلاقية في التعليم التقني: إدراج مقررات صلبة في الفلسفة الأخلاقية وأخلاقيات التكنولوجيا ضمن مناهج علوم الحاسوب والهندسة لترسيخ الوعي الأخلاقي لدى المطورين.
4. تأسيس مجالس وهيئات استشارية فلسفية: إنشاء لجان أخلاقية متخصصة ومستقلة في المؤسسات البحثية والشركات التقنية لمراجعة الآثار الوجودية والاجتماعية للبرمجيات قبل اعتمادها.

الخاتمة

تخلص هذه الدراسة النظرية المستفيضة إلى أن الفلسفة الأخلاقية ليست رفاهية فكرية أو ترفاً تنظيرياً في العصر الرقمي، بل هي الجدار الوقائي الأخير الذي يضمن الحفاظ على إنسانية الإنسان وكرامته. إن التطور المتسارع للذكاء الاصطناعي يحتم علينا عدم الانقذ المبهور خلف الإنجازات التقنية المجردة، بل يفرض ضرورة وضع التقنية في مكانها الطبيعي كـ "أداة خادمة للوجود الإنساني" وليست بديلة عنه أو هيمنة عليه. إن حماية المسؤولية والأخلاق تستدعي إعادة الاعتبار للفاعل البشري باعتباره المركز والغاية والأصل في كل تكليف أخلاقي وقانوني.

قائمة المصادر والمراجع

1. أحمد، سعيد حسام (2021). أخلاقيات التكنولوجيا: دراسة فلسفية نقدية في العصر الرقمي. دار الوفاء لدنيا الطباعة والنشر، الإسكندرية.
2. الجابري، محمد عابد (2018). الإنسان والتكنولوجيا: أبعاد الرؤية الفلسفية والمعاصرة. مركز دراسات الوحدة العربية، بيروت.
3. الشهري، عبد الله بن علي (2024). المسألة الأخلاقية في خوارزميات الذكاء الاصطناعي: منظور فلسفي إسلامي. مجلة الفكر المعاصر، المجلد 12، العدد 45، ص ص 112-140.
4. الخولي، يمني طريف (2019). فلسفة التكنولوجيا: من التنظير إلى الحوكمة الأخلاقية. دار المعارف، القاهرة.
5. مطر، أميرة حلمي (2019). الفلسفة الأخلاقية: نشأتها وتطورها. دار المعارف، القاهرة.
6. الهواري، خالد (2023). الإشكالات الوجودية للذكاء الاصطناعي: من الفاعل الإنساني إلى الفاعل الرقمي. مجلة أوراق فلسفية، العدد 68، ص ص 45-78.

7. قدورة، فاطمة محمد عبد السلام. (2026). أخلاقيات الذكاء الاصطناعي: مسؤولية الآلة ومسؤولية الإنسان في الفكر الإسلامي. مجلة دراسات الإنسان والمجتمع، 28.
8. ياسين، رينا غازي عبد الله. (2026). الذكاء الاصطناعي والمسؤولية الأخلاقية. مجلة كلية الأمة الجامعية للعلوم التربوية والنفسية.
9. بن زائد، نجاح أبو القاسم. (2025). فلسفة الذكاء الاصطناعي وأخلاقياته: قراءة فلسفية في السياق العربي. مجلة الأصالة، 6(11).
10. الحضيرى، كلثوم عثمان حسن. التحديات الأخلاقية للذكاء الاصطناعي. مجلة أبحاث، جامعة سبها.
11. بركة، سامي الكامل، وعمر، عبد الحكيم امحمد. (2026). الذكاء الاصطناعي بين الفلسفة والأخلاق وتحديات المستقبل. جامعة الزاوية.
12. عبد الباقي، صديقة. (2025). نظرة مقاصدية في تقنيات الذكاء الاصطناعي ومظاهر المسؤولية الأخلاقية في استخدامها. مجلة الحكمة للدراسات الفلسفية، 13(2)، 184-199.
13. هاجر، زيدان، وحنان، العابد. (2025). الذكاء الاصطناعي وحدوده الفلسفية والأخلاقية. رسالة ماجستير، جامعة الوادي.
14. (Coeckelbergh, M. (2024). أخلاقيات الذكاء الاصطناعي. ترجمة هبة عبد العزيز غانم. مؤسسة هنداوي.
15. سعيد، محمد طلعت. (2026). المعايير الأخلاقية والقانونية لاستخدامات الذكاء الاصطناعي. دار الأهرام القانونية.
16. ناصيف، إلياس. (2026). الذكاء الاصطناعي ومسؤولياته المدنية والجزائية.
17. Floridi, L., & Cowls, J. (2019). A Unified Framework of Five Principles for AI in Society. Harvard Data Science Review, 1(1).
18. Jobin, A., Ienca, M., & Vayena, E. (2019). The global landscape of AI ethics guidelines. Nature Machine Intelligence, 1, 389-399.
19. Floridi, L., et al. (2018). AI4People—An Ethical Framework for a Good AI Society. Minds and Machines, 28, 689-707.
20. Morley, J., Floridi, L., Kinsey, L., & Elhalal, A. (2019). From What to How: An Initial Review of Publicly Available AI Ethics Tools, Methods and Research. Science and Engineering Ethics.
21. Floridi, L., & Taddeo, M. (2016). What is Data Ethics? Philosophical Transactions of the Royal Society A, 374, 20160360.